



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2012

---

## **Multimodalität und multimodale Kompetenz. Text-Bild-Kommunikate im Social Web**

Müller, Christina Margrit

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-142708>

Journal Article

Published Version

Originally published at:

Müller, Christina Margrit (2012). Multimodalität und multimodale Kompetenz. Text-Bild-Kommunikate im Social Web. *Der Deutschunterricht*, 64(6):22-33.

# Multimodalität und multimodale Kompetenz

## Text-Bild-Kommunikate im Social Web

### BILDung

Bilder<sup>1</sup> spielen nicht nur im Social Web, sondern allgemein in der heutigen Kommunikation eine eminente Rolle. Hug (2011, 54) spricht von einem Visualisierungsschub, der sowohl quantitative als auch qualitative Züge trägt: Der qualitative Aspekt betrifft den Umgang mit Bildern sowie veränderte Sehpraxen, der quantitative ist auf die einfach gewordene Herstellung digitaler Bilder sowie deren massenhafte Nutzung zurückzuführen, nicht zuletzt in Foto- und Video-Communitys des Social Webs. *Pictorial turn* ist ein anderer geläufiger Terminus, der sowohl für die genannten aktuellen Veränderungen als auch historisch für essenzielle Wendepunkte wie beispielsweise die Entwicklung von Drucktechniken wie der Holzschnitt oder die Erfindung der Fotografie verwendet wird. „Gewöhnlich beschwören *pictorial turns* in irgendeiner Weise die Unterscheidung zwischen Worten und Bildern herauf, wobei das Wort mit Gesetz, Lesekompetenz und Elitenherrschaft, das Bild dagegen mit volkstümlichem Aberglauben, Unbildung und Ausschweifung assoziiert wird.“ (Mitchell 2009, 321, Hervorh. im Original) In einem solchen durch den Bedeutungszuwachs von Bildern evozierten kulturpessimistischen

Diskurs wird oftmals eine Verdrängung der geschriebenen Sprache befürchtet. Um solche Befürchtungen zu relativieren bedarf es wissenschaftlicher Forschung zur multimodalen Kommunikation (vgl. Klemm/Stöckl 2011, 7). Bei der sich gerade konstituierenden sogenannten ‚Bildlinguistik‘ handelt es sich um einen noch weitgehend unerforschten Teilbereich innerhalb der Text- und Medienlinguistik. Dass Bilder lange Zeit – mit Ausnahme einzelner Bereiche wie beispielsweise der Werbekommunikation – nicht in textlinguistische Analysen einbezogen wurden, wird auf den bisher logozentrischen<sup>2</sup> Blick bzw. die Bilderblindheit (vgl. Schmitz 2005, 197) der Sprachwissenschaft zurückgeführt.

Dies wiederum entspricht der Diskreditierung von Bildern in der Schule. Schmitz (2004, 73) fragt berechtigterweise: „Warum sollte Schule Bildung für Bilder vernachlässigen, wo visuelle Kommunikation doch unseren Alltag bestimmt?“ Denn gerade im Zeitalter des Social Webs werden sowohl von Produzenten als auch von Rezipienten „design- und bildbezogene Kompetenzen erwartet [...], die über eine sprachlich zu bestimmende Textkompetenz“ (Steinseifer 2011, 165) hinausgehen. Folglich wurden in den letzten Jahren immer mehr Stimmen laut, die fordern, dass Text-Bild-Kommuni-

(1) Dieser Arbeit liegt ein enger Bildbegriff zugrunde, d. h. als Bilder werden nur visuell wahrnehmbare Artefakte verstanden und nicht auch mentale Bilder. Solche materiellen Bilder gelten denn auch als prototypische Bilder (vgl. Stöckl 2004a, 91) und sie treten in den medialen Varianten ‚statisch‘ und ‚dynamisch‘ auf (vgl. Stöckl 2004b, 17). Der vorliegende Beitrag beschränkt sich auf statische Bilder.

(2) Logozentrismus kann dahingehend verstanden werden, dass der gesprochenen Sprache ein höherer Stellenwert als der geschriebenen zugestanden wird. Hier wird der Terminus jedoch in der Bedeutung verwendet, wie sie beispielsweise bei Stöckl (2004a, 6) zu finden ist, nämlich dass die Sprache sowohl medial schriftlich als auch mündlich als dem Bild überlegen gilt. Um den zweideutigen Terminus zu vermeiden, könnte man alternativ von „Linguozentrismus“ (Nöth 2000, 468) sprechen.

kation verstärkt Gegenstand des Schulunterrichts sein soll. So wird die Förderung visueller Kompetenz oder „Bildalphabetisierung“ (Doelker 2011, 10) im Schulunterricht gefordert. Mit dem Terminus der Alphabetisierung wird die bildungspolitische Forderung erhoben, „den etablierten Lesebegriff auch auf das Bild auszuweiten, also Bilderlesen als gleichwertige Kulturtechnik wie das Lesen von verbalen Texten zu positionieren und damit in den Alphabetisierungsauftrag der Schule einzu beziehen“ (Doelker 2011, 10). Diese Aussage basiert einerseits auf einem erweiterten Textbegriff, den wir, wie es Adamzik (2002, 175) dargelegt hat, im Gegensatz zu einer erweiterten Sichtweise, nicht brauchen. Andererseits beruht die Formulierung auf der Metapher des Lesens, deren Verwendung einhergeht „mit der übermäßigen Aufwertung einer Kultur des Textes und der Literarizität“ (Bachmann-Medick 2009, 349). Derlei metaphorische Übertragungen sollten vermieden werden, da sie auf einer logozentrischen Sichtweise beruhen und wiederum der Schrift Hegemonie attestieren. Ehlich (2005, 59) hingegen spricht analog zu der Aufgabe des Lesen-Lernens von derjenigen des Sehen-Lernens, welche uns vor neue didaktische Herausforderungen stellt. Eine Herausforderung ist es insofern, als die linguistische Forschung zur Text-Bild-Kommunikation erst in den Anfängen steckt; infolgedessen sind Text-Bild-Kommunikate auch noch kein Thema in der Ausbildung von Lehrpersonen und auch noch nicht in Lehrmitteln präsent (vgl. Ballstaedt 2005, 69). Und genau darin liegt gleichzeitig auch eine Chance: Das Thema bietet die Möglichkeit, „experimentell [...] didaktisches Neuland zu betreten“ (Ehlich 2005, 60) oder wie Schmitz es treffend formuliert: „Wo es noch keine Dogmen, eingefahrenen Lehrmeinungen, simplifiziert in den Schulalltag abgesunkenen Theoriefragmente und aus vierter Hand zitierenden Fibeln gibt, lässt sich frischer und unbefangener arbeiten. Entdeckendes Lernen entdeckt wirklich neu und muss nicht das finden, was Lehrmeister versteckt haben“ (Schmitz 2004, 72).

Der vorliegende Beitrag führt in die Forschung zur Text-Bild-Kommunikation mit speziellem Fokus auf das Social Web ein, gibt Anregungen für die Unterrichtsgestaltung und möchte vor allem dazu animieren, selber im Unterricht mit Text-Bild-Kommunikaten zu experimentieren und dazu die Möglichkeiten des Social Webs zu nutzen. Zunächst wird in Abschnitt 2 ein allgemeiner Überblick zur linguistischen Analyse von Text-Bild-Relationen gegeben, in Abschnitt 3 werde ich aufzeigen, inwiefern die genannten Typologien auf ein Beispiel des Social Webs angewendet werden können und in welcher Hinsicht Text-Bild-Relationen im Social Web sich von solchen in herkömmlichen multimodalen Kommunikaten unterscheiden. Dies ist deshalb von Bedeutung, weil intermodale Bezüge aufgrund von „Textsortengebundenheit und Kontextsensibilität“ (Stöckl 2006, 27) unterschiedlich geartet sein können, d. h. je nach Text-Bild-Sorte haben die verschiedenen Arten von Text-Bild-Relationen einen anderen Stellenwert. In Abschnitt 4 wird anhand von Studien zur Mediennutzung Jugendlicher dargelegt, dass Bilder im Alltag von Schülerinnen und Schülern eine zentrale Rolle spielen und in Abschnitt 5 schließlich wird ausgeführt, welche Kompetenzen in Bezug auf Text-Bild-Kommunikate im Unterricht erworben werden sollten.

### **Bildlinguistik: Analyse von Text-Bild-Kommunikaten**

Inzwischen ist man sich weitgehend einig darüber, dass in der Sprachwissenschaft „Sprache und mehr“ (Holly 2011, 160) untersucht werden muss. Hierzu gehören auch Text-Bild-Kommunikate, die Untersuchungsgegenstand der Bildlinguistik<sup>3</sup> sind. Darunter ist „die Betrachtung der Bezüge zwischen Sprache und Bild in Gesamttexten und die Nutzbarmachung linguistischer Konzepte, Modelle und Methoden für die Beforschung des in vorwiegend massenmediale Texte integrierten Bildes“ (Klemm/Stöckl 2011, 9) zu verstehen. Während früher Text-Bild-Kombinationen eher

(3) Als Einführung ins Themengebiet sei an dieser Stelle der gleichnamige Sammelband (Diekmannshenke et al. 2011) empfohlen.

### Text-Bild-Kommunikate im Social Web

Redundanz	im Bild wird dasselbe gezeigt, wie im Text gesagt wird
Dominanz	Hauptinformation im Text bzw. im Bild
Komplementarität	die Gesamtbedeutung ergibt sich aus der Text-Bild-Kombination
Diskrepanz	zusammenhangloses Nebeneinander von Text und Bild
Kontradiktion	Widerspruch zwischen Text und Bild

Tab. 1: Semantische Text-Bild-Relationen (Bearbeitung nach Luginbühl 2011, 258)

Ausnahmen waren, sind sie heutzutage der Normalfall; „der prototypische Text [ist] gegenwärtig kein rein sprachlicher mehr [...], sondern ein genuin multimodaler“ (Stöckl 2004a, 5). Unter Multimodalität<sup>4</sup> wiederum versteht man „Texte und kommunikative Handlungen, die mehrere verschiedene Zeichensysteme (Sprache, Bild, Ton) beinhalten“ (Stöckl 2011, 45). Der Terminus „Text“ erhält in diesem Zitat eine sehr weite Bedeutung, die von der Etymologie her (lat. *textus* = Gewebe, Verbindung, oder Zusammenhang) zwar ausgezeichnet passen würde, jedoch vom allgemeinen Sprachgebrauch, in welchem Text ausschließlich mit Schrift assoziiert wird, zu sehr abweicht (vgl. Adamzik 2002, 174) und deswegen wohl nicht konsensfähig werden kann. Darüber hinaus stellt Stöckl (2004a, 244) auch fest, dass der „Prototyp des Textes in logozentrischer Tradition primär sprachlich und zudem geschrieben oder gedruckt verfasst ist“; weitert man die Bedeutung auch auf andere als Modalitäten aus, wird jedoch meines Erachtens genau diese logozentrische Tradition fortgesetzt. Aus diesem Grunde folge ich Adamzik (2002, 174), die den Terminus ‚Kommunikat‘ anstelle von Text vorschlägt. Text-Bild-Kombinationen können als multimodale Kommunikate bezeichnet werden. Genau genommen müsste man natürlich bei Text-Bild-Kommunikaten von bimodalen Kommunikaten sprechen (vgl. Stöckl 2004a, 107), doch ist die Bezeichnung „multimodal“ als Hyperonym eben-

falls korrekt. Bisher standen im Zentrum der Forschung zur Multimodalität vor allem professionell erzeugte multimodale Kommunikate wie beispielsweise Werbeanzeigen oder journalistische Erzeugnisse (Print/TV). Multimodale Kommunikate von Kommunikationslaien sowie Alltagsbilder wurden bisher kaum untersucht (vgl. Schmitz 2005, 197); hierzu können größtenteils die multimodalen Kommunikationspraxen in Social-Networking- und Social-Sharing-Communitys gerechnet werden, die noch gänzlich unerforscht sind (vgl. Klemm/Stöckl 2011, 16).

Typologien von Text-Bild-Kommunikaten können aufgrund verschiedener Kriterien erstellt werden. Nöth (2000, 483) unterscheidet zwischen syntaktischen, pragmatischen und semantischen Gesichtspunkten. Unter syntaktisch versteht er die Anordnung von Text und Bild im Raum sowie deren Beziehungen. „Prinzipiell lassen sich deutliche graphische Trennungen der Sprach- und Bildteile eines Textes (Linearisierung) von einer eher räumlich integrierenden Inkorporation beider Codesysteme (Simultaneität) unterscheiden“ (Stöckl 2004a, 249). In pragmatischen Typologien werden „Formen der Bezugnahme, des Zeigens, vom Text zum Bild und vom Bild zum Text“ (Nöth 2000, 483) kategorisiert. Nöth (ebd., 485) spricht von „indexikalischen Verweisrelationen“, die verbal durch Deiktika oder nonverbal durch Bildelemente wie Pfeile realisiert werden können.

(4) Der Terminus wird außerdem verwendet für die „Verbindung von lautlichen mit mimischen, gestischen und anderen körperbezogenen Anteilen, und damit auditiver und visueller ‚Modalitäten‘ in der Face-to-face-Interaktion“ (Steinseifer 2011, 165). Zuweilen wurde und wird eine Differenzierung zwischen Multimodalität (mit unterschiedlichen Sinnen zu rezipierendes Kommunikat) und Multikodalität (aus verschiedenen Zeichenarten bestehendes Kommunikat) vorgenommen, doch hat sich diese Unterscheidung nicht durchgesetzt (vgl. Steinseifer 2011, 170).

Solche Relationen sind komplementär, da Text und Bild zwangsläufig aufeinander bezogen werden müssen. Komplementarität wiederum zählt Nöth zu den semantischen Aspekten. In semantischen Typologien „interessiert man sich v. a. für die Verteilung von Informationen auf Sprache und Bild sowie für das generelle semantische Verhältnis sprachlicher und visueller Botschaften zueinander“ (Stöckl 2004, 250; vgl. Tab. 1).

Die Unterscheidung zwischen Dominanz und Komplementarität kann nicht ganz trennscharf gezogen werden; bei der Dominanz sind die Informationen ungleich, bei der Komplementarität gleichmäßig auf die Modalitäten verteilt. Die Komplementarität ist denn auch die interessanteste Text-Bild-Relation und wurde verschiedentlich weiter differenziert.<sup>5</sup> Abschließend sei darauf hingewiesen, dass Bezüge zwischen Texten und Bildern nicht per se bestehen, sondern in der Interaktion des Rezipienten erzeugt werden (vgl. Bucher 2011, 146).

### Multimodale Kommunikation im Social Web

Multimodalität ist ein zentrales Merkmal der im Social Web stattfindenden Kommunikation. Insbesondere in Social-Sharing- aber auch in Social-Networking-Communitys spielen statische Bilder eine zentrale Rolle. Die bekannteste Social-Sharing-Community im Bereich statisches Bild ist Flickr, die über sechs Milliarden Fotos umfasst.<sup>6</sup> Die größte Fotosammlung im Social Web ist jedoch die Social-Networking-Community Facebook: Im Herbst 2011 umfasste sie bereits 140 Milliarden Bilder,<sup>7</sup> monatlich kommen mehrere Milliarden hinzu. Der Unterschied zwischen den beiden Arten von Communitys ist, dass die User andersartige Ziele verfolgen: Während bei Social-Sharing-Communitys der sogenannte *user generated content*, d. h. in die-

sem Falle die Bilder, im Mittelpunkt steht, geht es bei Social-Networking-Communitys vor allem um Kommunikation mit Familien, Freunden und Bekannten, welche nicht ausschließlich verbal, sondern auch über Bilder stattfindet. In beiden Arten von Communitys gibt es Möglichkeiten, die Bilder mit Texten zu versehen, seien dies Albumtitel, Bildtitel, Bildunterschrift, Tags (Schlagworte), Kommentare unterhalb oder innerhalb des Bildes. Die Texte können zum einen von demjenigen Nutzer, der das Bild hochgeladen hat, verfasst werden, zum anderen auch von allen anderen Usern, die berechtigt sind, das entsprechende Bild zu sehen und zu kommentieren. Terminologisch kann dies wie folgt ausdifferenziert werden: Die Teiltex-te, die vom Produzenten des Fotos verfasst wurden, bilden mit dem Bild zusammen ein multimodales Kommunikat. Sowohl die Teiltex-te als auch die Bilder können wiederum Kommunikat-teile genannt werden (vgl. Schneider/Stöckl 2011, 29). Die Kommentare, welche unterhalb des Bildes angebracht werden und sowohl Text- als auch Bildteile enthalten können, sowie die Notizen, d. h. Textteile, die innerhalb des Bildes platziert werden, können – sofern sie nicht vom Bild-Produzenten stammen – nicht zum multimodalen Kommunikat gerechnet werden. Die verschiedenen Kommentare können als einzelne kommunikative Akte bezeichnet werden, ein Terminus, den Jucker und Dürscheid im Hinblick auf Statusmeldungen auf Facebook oder Kommentare zu Online-Zeitungsartikeln vorgeschlagen haben. Sie verstehen darunter intentionale Kommunikation, die verbal oder auch non-verbal erfolgen kann. Mehrere kommunikative Akte, die nicht für sich allein stehen, sondern durch Bezüge eine Abfolge bilden, nennen Jucker und Dürscheid Sequenzen kommunikativer Akte.

Schauen wir uns hierzu in Abb. 1 ein konkretes Beispiel aus der Fotocommunity Flickr an. Das multimodale Kommunikat

(5) So untergliedert beispielsweise Doelker (2011, 23–24) die Kategorie der Komplementarität in die Teilkategorien kongruent, reziprok, komplementär, additiv, assoziativ und extensiv.

(6) Eine relativ neue Social-Sharing-Community ist Pinterest, in welcher die Nutzer thematisch geordnete Bildersammlungen mit einer kurzen Beschreibung anlegen können. Im Unterschied zu herkömmlichen Foto-Communitys geht es aber nicht um das Teilen selbst erstellter Bilder, sondern um Fundstücke aus dem Internet. Zwischen Mai 2011 und Januar 2012 ist die Zahl der Nutzer in Deutschland um 2956 % gestiegen, doch bleibt abzuwarten, ob die Community in Europa einen so großen Erfolg haben wird wie in den USA (vgl. Netlink 672).

(7) Netlink 673.



## Im Frühlingswald

Märzenbecher - vor einer Woche lag hier noch Schnee!

### Kommentare und Favoriten



**aerosmith6000000** (vor 2 Tagen)

Feiner POV....ich habe immer noch keine in der Natur gesehen, wahrscheinlich bin ich noch nicht wach genug für den Frühling...grins.  
Wünsche dir einen schönen Abend



**alfred.hausberger\_pro** (vor 2 Tagen)

Im Laub des Herbstes muss man sie zwar noch suchen, wenn man sie aber gefunden hat, leuchten sie umso schöner.



**baerchen57\_pro** (vor 2 Tagen)

ah, Du hast sie gefunden :-))

© www.flickr.com/photos/24973309@N04/6812943900/

Abb. 1: Beispiel für Bild, Bildtitel, Bildbeschreibung, Kommentare und Notiz<sup>8</sup>

besteht aus folgenden Elementen: einer Fotografie, dem Bildtitel „Im Frühlingswald“, der Bildbeschreibung „Märzenbecher – vor einer Woche lag hier noch Schnee!“ sowie den im Screenshot nicht zu sehenden Tags „Vorfrühling / Blumen / Märzenbecher / Wald / Natur / spring / flowers / Geretsried / Bavaria / Germany / Winter / Color Efex pro / Leica M9 / Voigtlander Nokton 35mm f1.4 SC“.

Bezüglich der in Abschnitt 2 genannten Typologien kann für die syntaktische Ebene angenommen werden, dass v. a. der Bildtitel und die -beschreibung aufgrund der Nähe zum Bild rezipiert werden. Auf pragma-

tischer Ebene kann das Lokaladverb *hier* genannt werden, welches auf das Foto bzw. auf den darauf abgebildeten Ort, aber auch auf den Bildtitel „Im Frühlingswald“ verweist. Weiter oben wurden solch indexikalische Verweisrelationen als zwingend komplementär charakterisiert; es zeigt sich an diesem Beispiel: „Die Komplementarität von Text und Bild kann auch eine Relation zwischen drei Konstituenten sein.“ (Nöth 2000, 484) Dass sogar mehr als drei Konstituenten beteiligt sein können, wird deutlich, wenn man die Tags einbezieht: Das Ortsadverb wird durch diese Angaben konkretisiert, es handelt sich um einen Frühlingswald im bayerischen Geretsried. Die Hauptinformation zur geografischen Lokalisierung liegt hier demnach dominant im Text. Bezüglich Redundanz zum Bild können Tags wie „Blumen“<sup>9</sup>, „Märzenbecher“, „Wald“ und „Natur“ genannt werden. Redundanz heißt zwar, dass auf einer Fotoseite dieselbe Information in Text und Bild zu finden ist, von der Funktion her sind diese Schlagworte jedoch keineswegs überflüssig. Sie dienen dazu, dass über die Suche, die nur verbal erfolgen kann, die Auffindbarkeit eines Fotos in der Social-Sharing-Community Flickr unter Milliarden anderer Fotos sichergestellt werden kann. Weitere vergebene Tags sagen nichts über die fotografierten Inhalte, sondern über die Entstehungsgeschichte des Bildes aus. So verweist „Leica M9“ auf die gebrauchte Kamera, „Voigtlander Nokton 35mm f1.4 SC“ auf das verwendete Objektiv und „Color Efex pro“ schließlich auf die benutzte Bildbearbeitungssoftware. Diskrepante oder kontradiktorische Relationen sind in diesem Beispiel keine zu finden.

Nun zu den kommunikativen Akten, die üblicherweise Reaktionen auf das gepostete Bild darstellen und in Form von Kommentaren und Notizen auftreten. Zunächst einmal ist festzuhalten, dass es sich um kurze Texte handelt. In einem Korpus von 1000 Flickr-Fotoseiten von 1000 verschiedenen Usern ist jeweils der erste abgegebene Kommentar durchschnittlich 55 Zeichen (inkl. Leerzeichen) lang, die Notizen

(8) Netlink 674; Eingabe unter [www.mediensprache.net](http://www.mediensprache.net). Aus Platzgründen sind nur die drei ersten Kommentare zu sehen.  
(9) Darüber hinaus kann eine Redundanz zwischen dem deutschsprachigen „Blumen“ und dem englischsprachigen „flowers“ registriert werden, sofern der Rezipient beide Sprachen beherrscht.

lediglich 40. In den drei Kommentaren in Abb. 1 lassen sich auch typische Merkmale der computervermittelten Kommunikation finden, so beispielsweise der nicht mit den sonst üblichen Asterisken gekennzeichnete Inflektiv *grins*, der zwinkernde Smiley ;-)), sprachökonomische Phänomene wie beispielsweise das Akronym *POV* (Point of view) sowie der Flüchtigkeitsfehler „war-scheinlich“.

Im Hinblick auf semantische Text-Bild-Relationen kann in den User-Reaktionen Redundanz in den Nomina „Natur“, „Frühling“ und „Laub des Herbstes“ gefunden werden. Komplementär zu den abgebildeten und in der Bildbeschreibung genannten Märzenbechern sind die Pronomina „keine“ und „sie“ (insgesamt viermal). Diskrepanz zum Bild ist schließlich der funktional der phatischen Kommunikation dienende Satz „Wünsche dir einen schönen Abend“ im ersten Kommentar. Auf der pragmatischen Ebene der Text-Bild-Relationen sind die Notizen von besonderer Bedeutung. Jeder Nutzer kann auf Bildern ein beliebig großes Rechteck markieren und einen Text von maximal 300 Zeichen hinzufügen. Dass Texte in Bilder hineingeschrieben werden, ist nichts Neues: Man denke beispielsweise an Werbeplakate oder Bildpostkarten. Neu ist allerdings, dass die Notizen nur eingeblendet werden, wenn man mit dem Cursor über die Bildfläche fährt, ansonsten verschwinden sie wieder. In Abb. 1 sind mit dem Rechteck die Märzenbecher markiert und mit dem Text „Vorschlag“ versehen. Es handelt sich bei diesem kommunikativen Akt um eine Crop-Empfehlung, d. h. einen Vorschlag zum Zuschneiden des Fotos. Dass dies eine gängige kommunikative Praxis in Fotocommunitys ist, ist andernorts detailliert nachzulesen (vgl. Müller 2012, 49–51). Notizen sind besonders dazu geeignet, ganz gezielt auf bestimmte Ausschnitte eines Bildes zu zeigen, was bedeutet, dass eine Text-Bild-Relation zustande kommt, bei der nur ein Teil des Bildes beteiligt ist. Vergleichbar ist dies mit Gesprächen vor Bildern, beispielsweise bei einem gemeinsamen Besuch in einer Kunstaustellung oder beim gemeinschaftlichen An-

sehen von Urlaubsfotos, bei welchen man per Zeigegeste auf einzelne Teile eines Bildes verweisen und diese entsprechend kommentieren kann.

Zusammenfassend kann für Fotoseiten in Social-Sharing-Communitys festgehalten werden: Größtenteils handelt es sich nicht um professionell erzeugte multimodale Kommunikate, weshalb die Text-Bild-Relationen in vielen Fällen weniger komplex ausfallen als zum Beispiel Kommunikate aus dem Bereich der Werbung, wo durch überraschende Effekte Aufmerksamkeit erzielt wird. Die Komplexität auf Fotoseiten besteht darin, dass Relationen häufig nicht nur zwischen zwei, sondern zwischen mehreren Elementen bestehen. Infolgedessen wird die Rezeption des Gesamtkommunikats sowie der darauf folgenden Sequenz kommunikativer Akte schwieriger, da einzelne Informationen an verschiedenen Stellen zusammengesucht werden müssen. Darüber hinaus sind es oftmals Bezüge zwischen Textteilen und Bildteilen, die der Rezipient herstellen muss. Insbesondere die Notizen-Kommunikation ist ein interessantes Beispiel dafür, ebenso wie für Relationen auf pragmatischer Ebene.

## Umgang Jugendlicher mit Bildern

„Medienwelten sind nicht nur Sprachwelten, sie sind vor allem Bilderwelten“ (Baacke 2007, 97). Dies gilt insbesondere auch für das Trägermedium<sup>10</sup> Computer, im Besonderen für solche, die durch das Internet vernetzt sind. Im World Wide Web, speziell im Social Web, spielen Bilder eine immer zentralere Rolle. Grund dafür ist die weite Verbreitung von digitalen Fotoapparaten sowie Internet-Breitbandanschlüssen. Immer mehr Internetnutzer verfügen über eine Flatrate, was dazu führt, dass viele – v.a. jugendliche – Nutzer permanent online sind. Schülerinnen und Schüler sind also tagtäglich mit solchen Bilderwelten im WWW konfrontiert. In Abb. 2 ist ersichtlich, wie sich die Mediennutzung von Jugendlichen in Deutschland (JIM-Studie) und der Deutschschweiz (JAMES-Studie)

(10) Medium wird hier im technologischen Sinne verwendet.

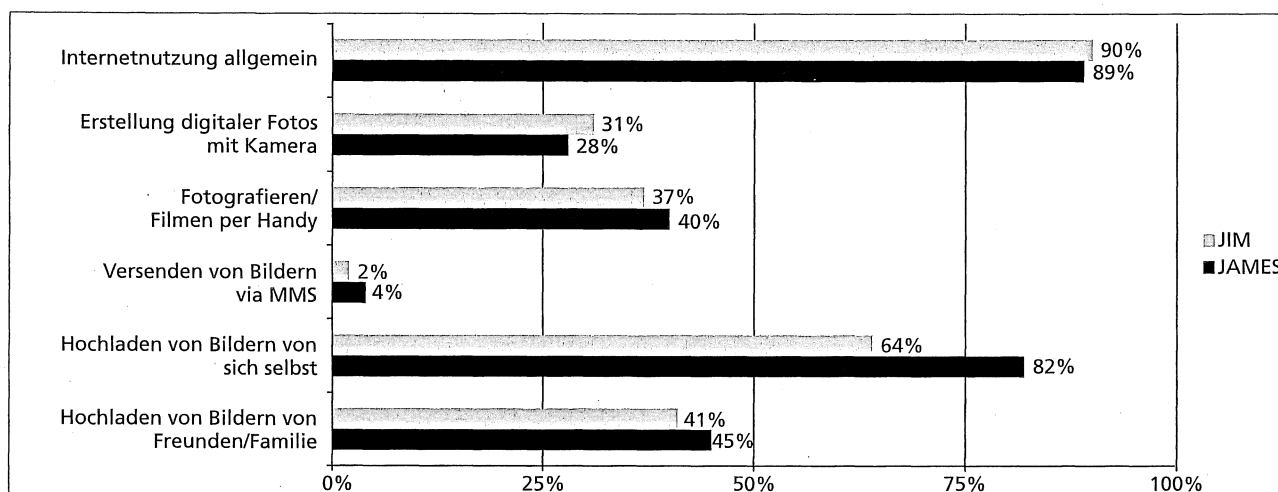


Abb. 2: Umgang Jugendlicher mit Bildern in Deutschland und der Deutschschweiz (Bearbeitung nach JIM-Studie 2010, 11, 44, 56; Willemse et al. 2010, 18, 32, 45)

in Bezug auf Bilder gestaltet.<sup>11</sup> Die Angaben zum Publizieren von Bildern beziehen sich darauf, ob im Jahr 2010 Fotos oder Videos hochgeladen wurden; die anderen Aktivitäten werden häufig, d. h. mehrmals pro Woche, ausgeführt.

Insgesamt kann konstatiert werden, dass die Werte in die beiden Länder sehr ähnlich sind. Die Unterschiede innerhalb der verschiedenen Sprachregionen der Schweiz (Deutschschweiz, Romandie, Tessin) fallen in einigen Punkten wie beispielsweise dem Upload von Bildern bedeutend größer aus als zwischen Deutschland und der Deutschschweiz (vgl. Willemse 2010, 29). In beiden Ländern ist der Versand von Bildern via MMS auffallend gering. Mit dem Aufkommen der Smartphones und Internet-Flatrates ist zu erwarten, dass die Nutzung von MMS weiter zurückgeht und dass mit dem Mobiltelefon aufgenommene Bilder in Zukunft eher direkt in Communitys des Social Webs hochgeladen werden und dann entweder via Handy oder Computer rezipiert werden.

Die Zahlen zeigen, dass das Internet und die digitale Fotografie im Alltag der Jugendlichen einen wichtigen Stellenwert haben. Im Rahmen einer Mediendidaktik ist es folglich notwendig, dass sich auch eine Bilddidaktik etabliert. Bilder wurden seit jeher für schulische Lernprozesse in unterschiedlichen Fächern eingesetzt, waren dabei aber häufig bloß Mittel zum Zweck. Im vorliegenden

Artikel wird jedoch das Bild als Unterrichtsgegenstand selbst ins Zentrum gerückt, wobei es ausschließlich um Bilder gehen soll, die in Kombination mit Sprache auftreten.

### Multimodale Kompetenz als Aufgabe des Deutschunterrichts

Warum sind Medien für den Deutschunterricht von Bedeutung? Frederking (2008, 89) argumentiert: „Der Deutschunterricht ist aufgrund der Medialität seiner unterrichtlichen Gegenstände Medienunterricht par excellence und deshalb in besonderer Weise für die Vermittlung von Medienkompetenz zuständig.“ Wenn man akzeptiert, dass der Mediendidaktik im Deutschunterricht ein wichtiger Stellenwert zukommt, dann muss man im gleichen Zuge auch zustimmen, dass neben Texten auch Bilder Unterrichtsgegenstand sein müssen. Die Frage ist berechtigt, warum Text-Bild-Kommunikate im Deutschunterricht behandelt werden sollten. Frederking (2008, 132) nennt zwei Gründe: 1) „weil viele visuelle Medienangebote auch sprachliche Anteile aufweisen“ und 2) „weil Verstehen eng mit sprachlichen Prozessen der Verbalisierung und der Kommunikation verknüpft ist. Bildrezeption (Wahrnehmen, Erschließen, Reflektieren) erfordert stets sprachlich-kommunikative Prozesse (z. B. Feststellen, Benennen, Paraphrasieren, Begründen)“. Selbstverständ-

(11) Die JIM-Studie 2011 ist zwar bereits publiziert, die JAMES-Studie 2011 jedoch noch nicht. Um die Zahlen vergleichen zu können, wurden deshalb beide Studien von 2010 verwendet. Die angegebenen Werte der JIM-Studie von 2010 unterscheiden sich außerdem von denjenigen von 2011 nur gering, jeweils maximal um drei Prozentpunkte.



lich sollen und müssen aber auch andere Fächer einbezogen werden. So ist es sinnvoll, wenn fächerübergreifende Lehr- und Lernzielbestimmungen formuliert werden (vgl. Ehlich 2005, 60). In erster Linie ist natürlich an den Kunstunterricht zu denken, aber auch weitere Fächer wie Geschichte, Sozialkunde und Religion können miteinbezogen werden (vgl. Schmitz 2004, 72). Darüber hinaus gibt es gerade bei digitalen Bildern wichtige ethische und rechtliche Fragen zu diskutieren,<sup>12</sup> wozu die entsprechenden Schulfächer einbezogen werden sollten.

Bislang hatten Text-Bild-Kombinationen im Deutschunterricht einen eher geringen Stellenwert, Comics sowie die Werbung bilden eine Ausnahme. Neben solchen von Profis erzeugten Kommunikaten müssen aber auch solche von Laien erstellte berücksichtigt werden. Diese spielen gegenwärtig im Social Web eine wichtige Rolle und sollten daher im Unterricht stärker thematisiert werden. Wie im Abschnitt 4 gezeigt, spielt multimodale Kommunikation im Social Web eine zentrale Rolle im Alltag vieler Jugendlicher. Daher ist es sinnvoll, im (Deutsch-)Unterricht den Lernenden die Möglichkeit zur Reflexion der eigenen kommunikativen Praktiken zu geben sowie ihre kommunikativen Kompetenzen im Bereich der Multimodalität zu verbessern. In diesem Zusammenhang ist von *visual literacy* die Rede. Was die Terminologie betrifft, so kann mit Hug (2011, 67) zur Kenntnis genommen werden, dass es „seit einiger Zeit Mode geworden [ist], neue Verständnisse von Literalität und Literacy zu generieren und auf verschiedenste Bereiche zu übertragen und in metaphorischer Weise anzuwenden.“ Die Bezeichnungen Literacy und Literalität sind jedoch meines Erachtens ungeeignet, da sie von lateinisch *littera* (= Buchstabe) abgeleitet werden. Aus diesem Grund ist der Ausdruck *visuelle Kompetenz* vorzuziehen.

Im Zusammenhang mit Text-Bild-Kommunikaten ist zu fragen, in welchem Verhältnis Medienkompetenz und visuelle Kompetenz

zueinander stehen. „Medienkompetenz ist im Kontext von Lese-, Schreib-, Informations- und Bildkompetenz zu sehen – sie schließt intermediale<sup>13</sup> Kompetenzen ein“ (Gapski 2009, 424). Zwar nennt Gapski auf der Ebene des Verbalen sowohl die Produktions- als auch die Rezeptionsseite, auf der Ebene des Bildes wird diese Differenzierung aber nicht vorgenommen. Auf die Frage nach der Relation von Medienkompetenz und visueller Kompetenz kann Folgendes geantwortet werden: Multimodale Kompetenz, die sich bei statischen Text-Bild-Kommunikaten aus einer Kombination aus visueller<sup>14</sup> und verbaler und bei Filmen auch auditiver Kompetenz zusammensetzt, ist ein Teilbereich der Medienkompetenz. Bei der multimodalen Kompetenz geht es um die Produktion und die Rezeption von multimodalen Kommunikaten, was „die Integration der verschiedenen Zeichenressourcen zu einem syntaktischen, semantischen und funktionalen Ganzen [erfordert]. Es sind diese kognitiven und textpraktischen Tätigkeiten, die man als ‚multimodale Kompetenz‘ fassen könnte“ (Stöckl 2011, 45). An selber Stelle führt Stöckl fünf Komponenten an, welche die multimodale Kompetenz umfasst, und zwar die Fähigkeiten

- 1) Sorten bzw. Typen von Bildern kategorisierend zu erkennen
- 2) dem Bild eine im Verwendungskontext relevante Bedeutung zuzuweisen
- 3) den Sprachtext im Abgleich mit der visuellen Botschaft zu verstehen
- 4) semantisierte Sprache und kontextualisiertes Bild zu integrieren
- 5) die Bildlichkeit der Sprache und der Textfläche bzw. des Schriftkörpers in den Prozess des Gesamtverstehens einzubeziehen

Betrachtet man die genannten Komponenten, so wird klar, dass 1) der visuellen Kompetenz zuzurechnen ist und damit aber ebenfalls eine Basis für die multimo-

(12) Zur Bildethik siehe Isermann/Knieper 2010, zu Urheberrechtsfragen und Creative Commons Weitzmann 2010.

(13) Gemeint sind in der Terminologie dieses Beitrags *intermodale Kompetenzen*, Stöckl (2011, 45) spricht von „multimodaler Kompetenz“.

(14) Visuell bezieht sich hier lediglich auf Bilder; selbstverständlich wird aber Sprache in der „medialen Variante“ (Stöckl 2011, 45) Schrift auch visuell rezipiert. Wir sehen daran, dass wir in der deutschen Sprache Bezeichnungen haben für „das Sehen/die Sprache betreffend“, aber (noch) keinen für „Bilder betreffend“.

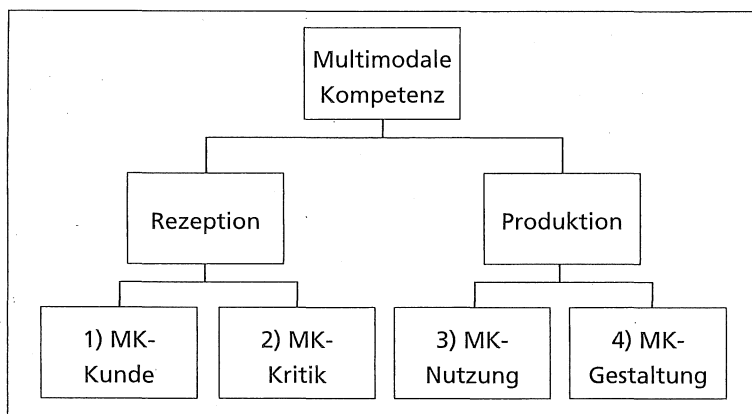


Abb. 3: Vier Teilbereiche multimodaler Kompetenz (MK = multimodales Kommunikat; Bearbeitung nach Baacke 2007, 98 f.)

dale Kompetenz darstellt, während 2) bis 5) sowohl visuelle als auch verbale Kompetenz erfordern und somit der multimodalen Kompetenz im engeren Sinne zugerechnet werden können. Nicht nur Stöckl, auch Ballstaedt (2005, 61–62) betont, dass die multimodalen Kompetenzen sowohl eine rezeptive als auch eine produktive Komponente voraussetzen, denn die beiden Bereiche sind eng miteinander verknüpft. Die in Abb. 3 aufgeführten Teilbereiche multimodaler Kompetenz können ebenfalls in Rezeption (1 und 2) und Produktion (3 und 4) multimodaler Kommunikate (im Folgenden abgekürzt: MK) unterteilt werden.

Das Ziel der MK-Kunde ist es, dass die Lernenden Kenntnisse über Funktions- und Wirkungsweisen von Text-Bild-Kommunikaten erwerben. Es geht zunächst darum, solche Kommunikate überhaupt verstehen zu können (vgl. Stöckls oben genannte Komponenten 2 und 3). Hierzu ist Wissen dazu notwendig, was grundsätzlich Unterschiede und Gemeinsamkeiten, bzw. Stärken und Schwächen von Text und Bild sind und in welcher Art und Weise sie zusammenwirken können (vgl. Abschnitt 2). Weiter ist es notwendig, Wissen im Bereich von rechtlichen und technischen Fragen bezüglich Bildern zu vermitteln, welches dann in der MK-Gestaltung benötigt wird. Ist einmal ein bestimmtes Grundwissen durch die MK-Kunde vorhanden, geht es darum, sich mit multimodalen Kommunikaten kritisch auseinanderzusetzen und diese zu reflektieren. Multimodale Kommunikate sollen auch bewertet und bezüglich ihrer Bedeutung eingeschätzt werden können, beispielsweise aus ethischer Perspektive.

Genauso wie die beiden Bereiche der MK-Kunde und der MK-Kritik eng miteinander verbunden sind, hängen MK-Nutzung und MK-Gestaltung zusammen. Bei der Nutzung geht es darum, bereits bestehende multimodale Kommunikate für eigene Anliegen zu nutzen, d. h. sinnvoll verwenden zu können, beispielsweise für schulische Aufgaben wie Referate. Die Lernenden sollen informationelle Autonomie erlangen, was bedeutet, dass es – noch bevor selbst multimodale Kommunikate gestaltet werden – wichtig ist, selbst Ressourcen (Bilder und Texte) beschaffen zu können: „Informationsautonom bedeutet, in der Lage zu sein, *selbst* auf die Informationsressourcen, die auf den Märkten oder in sozialen Beziehungen verfügbar sind, zugreifen und sie produktiv nutzen zu können.“ (Kuhlen 2004, 419, Hervorh. im Original) Sind entsprechende Ressourcen nicht vorhanden, können mit Rückgriff auf das in der MK-Kunde erworbene Wissen selber multimodale Kommunikate hergestellt werden. Voraussetzung dafür ist allerdings, dass die entsprechenden Fertigkeiten zur Gestaltung (v. a. die Bedienung entsprechender Software) vorhanden sind.

Nun geht es in diesem Beitrag speziell um multimodale Kompetenz im Social Web, d. h. es geht auch gleichzeitig um Digitalkompetenz, welche von Gapski (2009, 417) als Schlüsselkompetenz bezeichnet wird. Um eine solche Digitalkompetenz erreichen zu können, sollte zumindest ein Teil des Unterrichts zum Thema multimodale Kompetenz im Social Web in Form von E-Learning stattfinden. Durch das E-Learning verändert sich auch die Rolle der Lehrpersonen, die darin besteht, die Lernprozesse hauptsächlich zu coachen; Röll (2010, 203) spricht im Zusammenhang mit dem Social Web sehr treffend von einer „Pädagogik der Navigation“.

Abschließend seien noch einige Anregungen für die konkrete Unterrichtsgestaltung gegeben: Gerade im Social Web bietet es sich an, mit anderen Klassen (ev. aus anderen Schulen) gemeinsam ein Projekt durchzuführen. Sinnvollerweise arbeitet man dabei mit einem eigens dafür gewählten Tag, über welches in einer Community jeweils die zugehörigen Kommunikate gefunden werden können. Im Unterricht sollte mit

der Ebene der Rezeption (MK-Kunde und -Kritik) begonnen und danach zur Ebene der Produktion (MK-Nutzung und -Gestaltung) übergegangen werden. Gerade im Bereich der Rezeption bietet es sich im Zusammenhang mit dem Social Web an, mit älteren Lernenden das Themengebiet zu erarbeiten. Diese von den jüngeren Schülerinnen und Schülern als kompetent wahrgenommene Peers können im Anschluss daran diese unterrichten und multimodale Kompetenz vermitteln. Zunächst ist es zweckmäßig, wenn die Lernenden Text-Bild-Kommunikate im Social Web sammeln, analysieren und kategorisieren, also Text-Bild-Sorten definieren. Zum Sammeln und Archivieren von Beispielen sei das kostenlose Archivierungstool *Evernote* empfohlen. Bevor die Schülerinnen und Schüler selbst Text-Bild-Kommunikate erstellen, bei denen sie bereits bestehende Bilder verwenden, müssen rechtliche Fragen thematisiert werden, insbesondere relevant ist „die Frage nach dem Copyright [...] sowie den Rechten zur Bearbeitung, Verwertung und eigenen Nutzung behandeln“ (Geldhauser 2006, 74). Für den Unterricht geeignet sind Creative-Commons-Bilder, die bearbeitet werden dürfen.<sup>15</sup> So können Manipulationen an Text oder Bild vorgenommen werden, um daraufhin zu analysieren, inwiefern sich die Bedeutung des Gesamtkommunikats verändert. Schließlich sollen die Lernenden auch eigene Text-Bild-Kommunikate erstellen, die möglichst einen praktischen Nutzen erfüllen, beispielsweise für die Website der Schule oder eine Schülerzeitung im Web. Gerade in Fotocommunitys, welche eine Notizfunktion aufweisen, ist auch das Erstellen von Fotoromanen<sup>16</sup> (vgl. Reuen 2005) reizvoll: Sprechblasen oder auch Dialoge (vgl. Müller 2010, 242–250) können über Notizen erstellt werden. Für die Produktion steht eine Vielzahl an Open-Source-Software zur Verfügung: „Kostenlos verfügbare Software ermöglicht auch weniger finanzkräftigen Schülern legal zu Hause an den erstellten Materialien arbeiten und das Erlernte vertiefen zu können.“

(Geldhauser 2006, 76) Grundsätzlich sollen die erstellten Produkte der Öffentlichkeit zugänglich gemacht werden, wozu das Internet sehr geeignet ist. Dies bietet nicht zuletzt die Chance, den Eltern oder anderen Interessierten einen Einblick in die Arbeit des Deutschunterrichts zu gewähren.

## Ausblick und Fazit

Im vorliegenden Beitrag ging es lediglich um Text-Bild-Kommunikate. Selbstverständlich müssen aber auch die zentralen Modalitäten Geräusch und Musik sowie das dynamische Bild und die gesprochene Sprache einbezogen werden, damit die multimodale Kompetenz umfassend gefördert werden kann. Ferner sollen in diesem Zusammenhang auch Typografie und Layout thematisiert werden.

Es wurde im Beitrag aufgezeigt, welchen wichtigen Stellenwert multimodale Kommunikate heutzutage in unserer Gesellschaft haben. Wer über keine multimodale Kompetenz verfügt, der kann nicht problemlos Informationen rezipieren, da diese heutzutage oftmals multimodal präsentiert werden. Aus diesem Grund ist multimodale Kompetenz eine Schlüsselkompetenz, die im Unterricht vermittelt werden muss. Da Multimodalität ein typisches Merkmal des Social Webs ist, bietet es sich an, multimodale Kompetenz im WWW zu fördern, doch soll dies selbstverständlich auch offline geschehen. Dabei spielen sowohl die Rezeption als auch die Produktion von multimodalen Kommunikaten eine wichtige Rolle. Bei der Klassifikation von Text-Bild-Kommunikaten sollen sowohl solche aus dem Social Web als auch solche in Printmedien berücksichtigt werden. Wenn die Schülerinnen und Schüler durch die Rezeption und Produktion von Text-Bild-Kommunikaten das Wissen erworben haben, welches die spezifischen Schwächen und Stärken der beiden Modalitäten sind und wie sie zusammenwirken, ist ein wichtiges Ziel in diesem Bereich der BILDung erreicht. ■

(15) Unter Netlink 676 können in einer Metasuche solche Bilder in verschiedenen Communitys gefunden werden.

(16) Als Anregung kann das Beispiel unter Netlink 675 dienen.

## Literatur

- Adamzik, Kirsten (2002): Zum Problem des Textbegriffs. Rückblick auf eine Diskussion. In: Fix, Ulla/Adamzik, Kirsten/Antos, Gerd/Klemm, Michael (Hg.): Brauchen wir einen neuen Textbegriff? Antworten auf eine Preisfrage. Frankfurt am Main (= Forum angewandte Linguistik, 40), 163–182.
- Baacke, Dieter (2007): Medienpädagogik. Tübingen (= Grundlagen der Medienkommunikation, 1).
- Bachmann-Medick, Doris (2009): Cultural turns. Neuorientierungen in den Kulturwissenschaften. 3., neu bearb. Aufl. Reinbek bei Hamburg.
- Ballstaedt, Steffen-Peter (2005): Text-Bild-Kompositionen im Unterrichtsmaterial. In: Der Deutschunterricht 4, 61–70.
- Bucher, Hans-Jürgen (2011): Multimodales Verstehen oder Rezeption als Interaktion. Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): Bildlinguistik. Theorien – Methoden – Fallbeispiele. Berlin (= Philologische Studien und Quellen, 228), 123–156.
- Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.) (2011): Bildlinguistik. Theorien – Methoden – Fallbeispiele. Berlin (= Philologische Studien und Quellen, 228).
- Doelker, Christian (2011): Visuelle Kompetenz. Grundzüge der Bildsemantik. In: Hug, Theo/Kriwak, Andreas (Hg.): Visuelle Kompetenz. Beiträge des interfakultären Forums. Innsbruck, 9–27.
- Ehlich, Konrad (2005): Sind Bilder Texte? In: Der Deutschunterricht 4, 51–60.
- Frederking, Volker/Krommer, Axel/Maiwald, Klaus (2008): Mediendidaktik Deutsch. Eine Einführung. Berlin (= Grundlagen der Germanistik, 44).
- Gapski, Harald (2009): Medienkompetenz als gesellschaftliche Schlüsselkompetenz. In: Robertson-von Trotha, Caroline Y. (Hg.): Schlüsselqualifikationen für Studium, Beruf und Gesellschaft. Technische Universitäten im Kontext der Kompetenzdiskussion. Karlsruhe (= Schriftenreihe des Zentrums für Angewandte Kulturwissenschaft und Studium Generale der Universität Karlsruhe, 14), 409–426.
- Geldhauser, Steve (2006): Di(gi)daktik. Didaktische Potenziale von Digitalfotografie und Text. In: Holzbrecher, Alfred (Hg.): Foto + Text. Handbuch für die Bildungsarbeit. Wiesbaden, 73–82.
- Holly, Werner (2011): Medien, Kommunikationsformen, Textsortenfamilien. In: Habscheid, Stephan (Hg.): Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation. Berlin, 144–163.
- Hug, Theo (2011): Visuelle Kompetenz, Medienkompetenz und „New Literacies“. Konzeptionelle Überlegungen in einer pluralen Diskurslandschaft. In: Hug, Theo/Kriwak, Andreas (Hg.): Visuelle Kompetenz. Beiträge des interfakultären Forums. Innsbruck, 54–74.
- Isermann, Holger/Knieper, Thomas (2010): Bildethik. In: Schicha, Christian/Brosda, Carsten (Hg.): Handbuch Medienethik. Wiesbaden, 304–317.
- JIM (2011): Jugend, Information, (Multi-)Media. Basisstudie zum Medienumgang 12- bis 19-Jähriger in Deutschland, hg. vom Medienpädagogischen Forschungsverbund Südwest. <<http://www.mpfs.de/fileadmin/JIM-pdf11/JIM2011.pdf>>.
- Jucker, Andreas H./Dürscheid, Christa (eingereicht): The linguistics of keyboard-to-screen communication. A new terminological framework. In: Linguistik online.
- Klemm, Michael/Stöckl, Hartmut (2011): Bildlinguistik – Standortbestimmung, Überblick, Forschungsdesiderate. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): Bildlinguistik. Theorien – Methoden – Fallbeispiele. Berlin (= Philologische Studien und Quellen, 228), 7–18.
- Kuhlen, Rainer (2004): Informationsethik. Umgang mit Wissen und Information in elektronischen Räumen. Konstanz.
- Luginbühl, Martin (2011): Vom kommentierten Realfilm zum multimodalen Komplex. Sprache-Bild-Beziehungen in Fernsehnachrichten im diachronen und internationalen Vergleich. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): Bildlinguistik. Theorien – Methoden – Fallbeispiele. Berlin (= Philologische Studien und Quellen, 228), 257–276.
- Mitchell, W.J.T. (2009): Vier Grundbegriffe der Bildwissenschaft. In: Sachs-Hombach, Klaus (Hg.): Bildtheorien. Anthropologische und kulturelle Grundlagen des Visualistic Turn. Frankfurt am Main, 319–327.
- Müller, Christina Margrit (2010): SpielCommunitySpiel. Notizen auf der Online-Plattform Flickr. In: Stöckl, Hartmut (Hg.): Mediale Transkodierungen. Metamorphosen zwischen Sprache, Bild und Ton. Heidelberg (= Wissenschaft und Kunst, 17), 231–252.
- Müller, Christina Margrit (2012): Kommunikation im Bild: Notizen in Fotocommunitys. In: Siever, Torsten/Schlobinski, Peter (Hg.): Web 2.0. Innovationen zum III. Workshop zur linguistischen Internetforschung (= Sprache – Medien – Innovationen, 3), 39–62.
- Nöth, Winfried (2000): Handbuch der Semiotik. 2., vollst. neu bearb. und erw. Aufl. Stuttgart.
- Reuen, Sascha (2005): Fotoromane im Deutschunterricht. In: Der Deutschunterricht 4, 77–84.
- Röll, Franz Josef (2010): Web 2.0 als pädagogische Herausforderung. In: Herzig, Bardo/Meister, Dorothee M./Moser, Heinz/Niesyto, Horst (Hg.): Jahrbuch Medienpädagogik 8. Medienkompetenz und Web 2.0. Wiesbaden, 201–220.

- Schmitz, Ulrich (2004): Schrift und Bild im öffentlichen Raum. In: Mitteilungen des Deutschen Germanistenverbandes, Jg. 51 (1), 58–74.
- Schmitz, Ulrich (2005): Blind für Bilder. Warum sogar Sprachwissenschaftler auch Bilder betrachten müssen. In: Haueis, Eduard (Hg.): Paradigms lost. Duisburg (= Osnabrücker Beiträge zur Sprachtheorie, 69), 187–227.
- Schneider, Jan Georg/Stöckl, Hartmut (2011): Medientheorien und Multimodalität: Zur Einführung. In: Schneider, Jan Georg/Stöckl, Hartmut (Hg.): Medientheorien und Multimodalität. Ein TV-Werbespot – Sieben methodische Beschreibungsansätze. Köln, 10–38.
- Steinseifer, Martin (2011): Die Typologisierung multimodaler Kommunikationsangebote. Am Beispiel der visuellen Aspekte seitenbasierter Dokumente. In: Habscheid, Stephan (Hg.): Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation. Berlin, 164–189.
- Stöckl, Hartmut (2004a): Die Sprache im Bild – Das Bild in der Sprache. Zur Verknüpfung von Sprache und Bild im massenmedialen Text. Konzepte, Theorien, Analysemethoden. Berlin (= Linguistik – Impulse & Tendenzen, 3).
- Stöckl, Hartmut (2004b): Typographie: Gewand und Körper des Textes – Linguistische Überlegungen zu typographischer Gestaltung. In: Zeitschrift für Angewandte Linguistik (ZfAL) 41, 5–48.
- Stöckl, Hartmut (2006): Zeichen, Text und Sinn. Theorie und Praxis der multimodalen Textanalyse. In: Eckkrammer, Eva Martha/Held, Gudrun (Hg.): Textsemiotik. Studien zu multimodalen Texten. Frankfurt am Main, New York (= Sprache im Kontext, 23), 11–36.
- Stöckl, Hartmut (2011): Sprache-Bild-Texte lesen. Bausteine zur Methodik einer Grundkompetenz. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): Bildlinguistik. Theorien – Methoden – Fallbeispiele. Berlin (= Philologische Studien und Quellen, 228), 45–70.
- Weitzmann, John Hendrik (2010): Im Lizenzbaumarkt. Creative Commons als alternatives Modell. In: Heinrich-Böll-Stiftung in Zusammenarbeit mit iRights.info (Hg.): Copy. Right.Now! Plädoyers für ein zukunftstaugliches Urheberrecht. Berlin (= Schriften zu Bildung und Kultur, 4), 73–77.
- Willemsen, Isabel/Waller, Gregor/Süss, Daniel (2010): JAME Jugend. Aktivitäten. Medien. Erhebung Schweiz. Zürcher Hochschule für Angewandte Wissenschaften. Online verfügbar unter [http://www.psychologie.zhaw.ch/fileadmin/user\\_upload/psychologie/Downloads/Forschung/James/Ergebnisbericht\\_JAMES\\_2010\\_de.pdf](http://www.psychologie.zhaw.ch/fileadmin/user_upload/psychologie/Downloads/Forschung/James/Ergebnisbericht_JAMES_2010_de.pdf), zuletzt geprüft am 28.3.2012.